*J. Serb. Chem. Soc. 71* (7) 745–769 (2006)
*JSCS – 3468*

UDC 665.61+543.544.3:547.538+547.597
*Original scientific paper*

# Multivariate statistical methods applied to interpretation of saturated biomarkers (Velebit oil field, SE Pannonian Basin, Serbia)

TATJANA ŠOLEVIĆ[1#], KSENIJA STOJANOVIĆ[1,2], BRANIMIR JOVANČIĆEVIĆ[1,3*#], GORICA MANDIĆ[4], JAN SCHWARZBAUER[5] AND DRAGOMIR VITOROVIĆ[1,3#]

*[1]Center of Chemistry, ICTM, Njegoševa 12, 11000 Belgrade, [2]Department of Medicine, University of Belgrade, Višegradska 26, 11000 Belgrade, [3]Department of Chemistry, University of Belgrade, PO Box 158, 11001 Belgrade, [4]NIS – Naftagas, Narodnog fronta 12, 21000 Novi Sad, Serbia and [5]Institute of Geology and Geochemistry of Petroleum and Coal, Aachen Univeristy, Lochnerstrasse 4-20, D-52056 Aachen, Germany*
*(E-mail: bjovanci@chem.bg.ac.yu)*

(Received 5 October 2005)

*Abstract*: Twenty-five crude oils originating from the Velebit oil field (SE Pannonian Basin), the most important oil field in Serbia, were investigated. Saturated biomarkers (*n*-alkanes, isoprenoids, steranes and triterpanes) were analyzed by gas chromatography-mass spectrometry (GC-MS). Based on the distribution and abundance of these compounds, a large number of source and maturation parameters were calculated, particularly those most often used in correlation studies of oils. The examined samples were classified according to their origin and level of thermal maturity using factor, cluster and discriminant analyses. According to the source and maturation parameters, combined factor and cluster analyses using the Ward method enabled the categorization of the investigated oils into three groups. The cluster Ward analysis was shown to be of greater susceptibility and reliability. However, in addition to the two aforementioned methods, K-Means cluster analysis and discriminant analysis were shown to be necessary for a more precise and detailed categorization in the case of a large number of samples in one group. Consequently, it was concluded that factor and cluster K-Means and Ward analyses can generally be used for the interpretation of saturated biomarkers in correlation studies of oils, but the observed results have to be checked, *i.e.*, confirmed by discriminant analysis.

*Keywords*: crude oils classification, steranes, terpanes, source and maturation parameters, factor analysis, cluster K-Means and Ward analyses, discriminant analysis.

## INTRODUCTION

With the development of various analytical instrumental methods, a large number of compounds in crude oils have been discovered. Owing to organic-geo-

* Corresponding author.
\# Serbian Chemical Society active member.

doi: 10.2298/JSC0607745S

chemical research, their genesis and further transformations in the geosphere were elucidated, as well as the effects of the lithological composition of source and reservoir rocks on these processes.[1–3] Consequently, a large number of organic-geochemical parameters has been proposed to explain their origin, sedimentation environment, level of thermal maturity, length of migration and degree of microbiological degradation in reservoir rocks. However, it is well known that parameters influenced by only one organic-geochemical process are very rare, or even do not exist.[2,4,5] For example, the distribution and abundance of the very well examined saturated biomarkers in oils, particularly of *n*-alkanes and isoprenoid aliphatic alkanes, are influenced by the origin, depositional environment, thermal maturity and biodegradation. Furthermore, it is well known that in hopane and sterane isomerization reactions (moretane $\rightarrow$ hopane, 22*R* $\rightarrow$ 22*S* hopanes, 14α(H)17α(H) $\rightarrow$ 14β(H)17β(H) steranes, 20*R* $\rightarrow$ 20*S* steranes) the equilibria were established before the end of the catagenetic changes of the organic matter (*R*o = ~ 0.65 % and *R*o = ~ 1.00 %, respectively).[3] Finally, parameters calculated from distributions and abundances of saturated biomarkers are influenced by the lithological composition of the source and reservoir rocks, *i.e.*, their constituting minerals do not have the same catalytic effect on reactions resulting in their structural and stereochemical transformations.[3,6] Therefore, a reliable correlation of oils requires as many parameters as possible. On the other hand, a large number of parameters calculated for a large number of samples make the interpretation of the results difficult. In such cases, multivariate statistical methods, such as factor, cluster and discriminant analyses, become specially useful in correlation studies of oils.[7–11]

Twenty-five crude oil samples originating from the Velebit oil field, the most important oil field in Serbia, were investigated in this study. Saturated biomarkers (*n*-alkanes, isoprenoids, steranes and triterpanes) were analysed by gas chromatography-mass spectrometry (GC-MS). Considering the fact that all the samples were taken from the same oil field, in which all oil wells were at approximately the same depths of steady pressure and temperature, the values of the source and maturation parameters based on saturated hydrocarbons, except for one oil, were placed in a relatively narrow range. Because of this, an effort was made to make a better, more detailed classification of the samples based on source and maturation parameters, using multivariate statistical methods, comprising factor, cluster (K-Means and hierarchical, Ward) and discriminant analyses, which offer the possibility of examining a large number of calculated parameters simultaneously. Moreover, different statistical methods were used with the aim of determining which one of them would offer the most precise classification and thus could be regarded as the most applicable for the interpretation of saturated biomarkers in correlation studies of oils.

By combining the three statistical methods, the investigated oils were classified into three groups according to their source and maturation parameters. Furthermore, the samples from one group were additionally categorized into five sub-

groups according to their origin, and into three subgroups according to their thermal maturity.

## EXPERIMENTAL

*Samples and methods*

*Samples*. The oil-gas field Velebit is located in the southeastern part of the Pannonian Basin. This basin, with a surface area of approximately 260 000 km$^2$, belongs to the category of hyperthermal basins, characterized by high geothermic gradients, in some parts up to 70 ºC/km. Such a high geothermic gradient provided satisfactory heating even of Pliocene sediments, which can be considered as a bedrock regardless of the geological age.[12] The Pannonian Basin is a complex geological structure. It is not a uniform sedimentary basin. Intense tectonic activity, followed by volcanism of variable intensity, led during the Neogene and Anthropogene to the formation of a number of tectonic depressions, characterized by specific geological and geotectonic development.[13,14] The oil window in this basin is considered to begin at a depth of 2–3 km and to end at 3.5 to 5 km (corresponding to vitrinite reflection of *R*o = 0.60 – 1.30 %). Hence, most of the Middle Miocene and Lower Pannon sediments already passed through the oil window phase.[15,16]

The most important oil and gas deposits in the Serbian part of the Pannonian Basin have been discovered in the Banat depression. Its major part is located on the territory of Vojvodina (north of the rivers Sava and Danube), while its smaller, southern part, is located in the vicinity of the city of Požarevac (south of the Sava and the Danube), forming a separate smaller depression, the Drmno depression (Fig. 1).

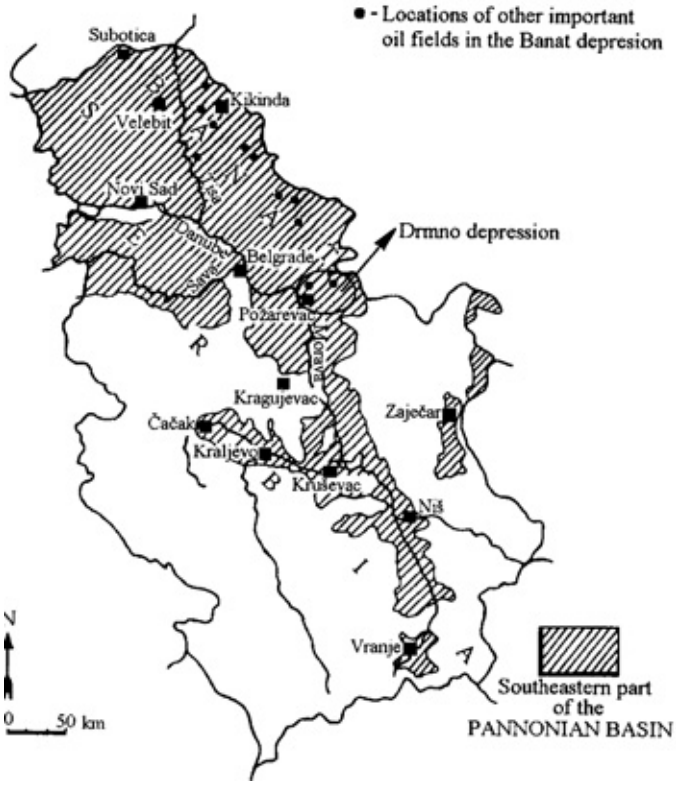


Fig. 1. Banat depression with locations of the most important oil fields.

The twenty-five crude oil samples investigated in this study originated from the largest oil-gas field Velebit, located in the part of the Banat depression belonging to the territory of the Province of Vojvodina (Fig. 1). During the year 2004, in this oil field, 138 oil wells operated, producing 230 000 t of oil and 14 800 000 $m^3$ of gas. Oil-gas field Velebit belongs to Pannonian lower pont complex. This complex is composed of limestone, siltstone and sandstone, and thus has favourable collector characteristics.

TABLE I. Oils from the Velebit oil field

| No | Depth/m | Lithology | Stratigraphy | Temperature in the reservoir/°C | Pressure in the reservoir/bar |
|---|---|---|---|---|---|
| 1 | 749.00–751.00 | Sandstone | Lower Pont | 68.0 | 74.87 |
| 2 | 753.50–757.50 | Sandstone | Lower Pont | 68.0 | 75.45 |
| 3 | 751.00–753.00 | Sandstone | Lower Pont | 68.0 | 74.08 |
| 4 | 759.00–762.00 | Sandstone | Lower Pont | 69.5 | 73.88 |
| 5 | 755.00–758.00 | Sandstone | Lower Pont | 69.0 | 71.36 |
| 6 | 758.00–762.00 | Sandstone | Lower Pont | 69.0 | 70.68 |
| 7 | 755.00–758.00 | Sandstone | Lower Pont | 69.0 | 72.52 |
| 8 | 758.00–763.00 | Clayey sandstone | Lower Pont | 69.5 | 72.44 |
| 9 | 751.00–752.00 | Sandstone | Lower Pont | 68.0 | 70.53 |
| 10 | 748.50–750.50 | Sandstone | Lower Pont | 68.0 | 70.66 |
| 11 | 759.50–761.50; 755.50–757.50 | Sandstone | Lower Pont | / | / |
| 12 | 760.00–761.00 | Sandstone | Lower Pont | 68.5 | 73.27 |
| 13 | 756.50–758.00 | Sandstone | Lower Pont | 68.0 | 71.85 |
| 14 | 751.00–753.73 | Sandstone | Lower Pont | 68.0 | 72.21 |
| 15 | 774.00–776.00 | Sandstone | Lower Pont | / | / |
| 16 | 751.00–754.00 | Clayey sandstone | Lower Pont | / | / |
| 17 | 746.00–747.00 | Sandstone | Lower Pont | / | / |
| 18 | 768.00–770.00 | Sandstone | Lower Pont | 70.0 | 70.47 |
| 19 | 783.50–785.00 | Sandstone | Lower Pont | / | / |
| 20 | 764.00–766.00 | Sandstone | Lower Pont | 70.0 | 68.42 |
| 21 | 758.00–765.00 | Sandstone | Lower Pont | 69.5 | 71.43 |
| 22 | 777.00–779.00 | Sandstone | Lower Pont | / | / |
| 23 | 770.00–771.00 | Sandstone | Lower Pont | / | / |
| 24 | 755.00–758.00 | Sand | Lower Pont | / | / |
| 25 | 750.00–752.00 | Sandstone | Lower Pont | 68.0 | 69.11 |

Lower Pont (Miocene)

This complex is regionally widespread. At some points it is up to 1000 m wide. In addition to the oil-gas field Velebit, nine more oil deposits were discovered in the Pannonian lower pont complex,[12,17] but the locations of the corresponding source rocks have so far not been determined. The list of the investigated samples, including data on depths, temperature, lithology and age of their res-

ervoir rocks, is given in Table I. The locations of the corresponding oil wells in the oil field Velebit are shown in Fig. 2.

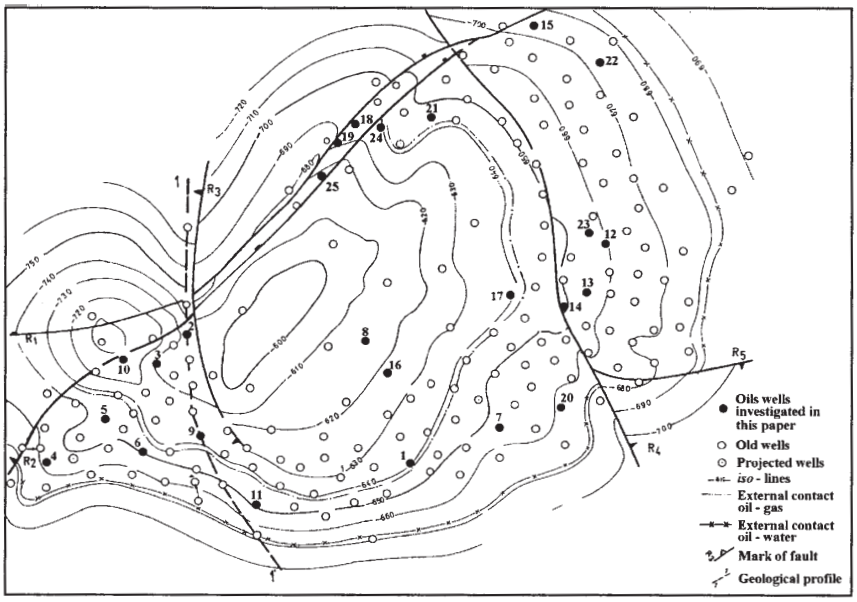


Fig. 2. Locations of wells in the oil-gas field Velebit.

*Methods*. The saturated hydrocarbons were isolated using a procedure consisting of preliminary elimination of water (by a mixture of *n*-heptane and toluene, 7:3), asphaltenes (by *n*-heptane), and NSO compounds (by *n*-hexane; Soxhlet), followed by column chromatography ($\gamma$-$Al_2O_3$; 80:1) using *n*-hexane as the eluent.

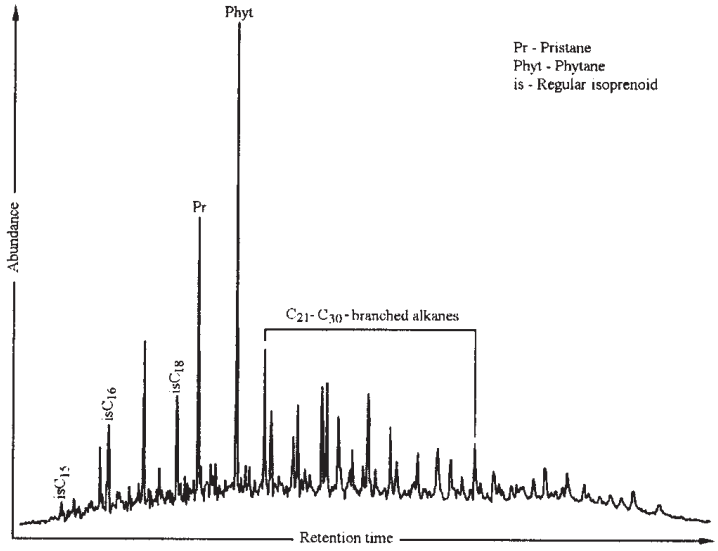


Fig. 3. GC-MS fragmentogram of alkanes (*m/z* 71) isolated from sample Velebit 3, typical for the investigated oils.

The saturated hydrocarbon fractions were analysed by the gas chromatographic-mass spectrometric technique (GC-MSD), using the single ion monitoring (SIM) method. A Carlo Erba Instru-

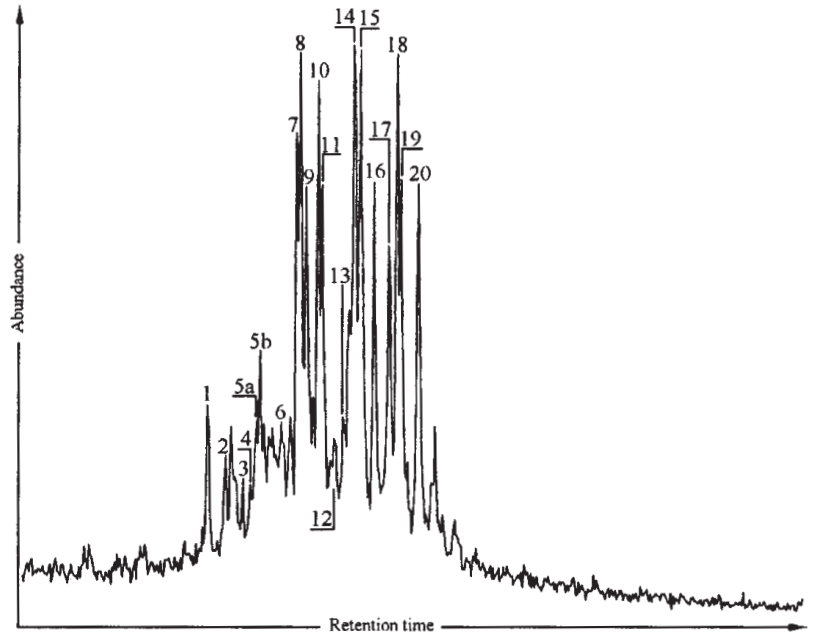


Fig. 4. GC-MS fragmentogram of steranes (*m/z* 217) isolated from sample Velebit 3, typical for the investigated oils.

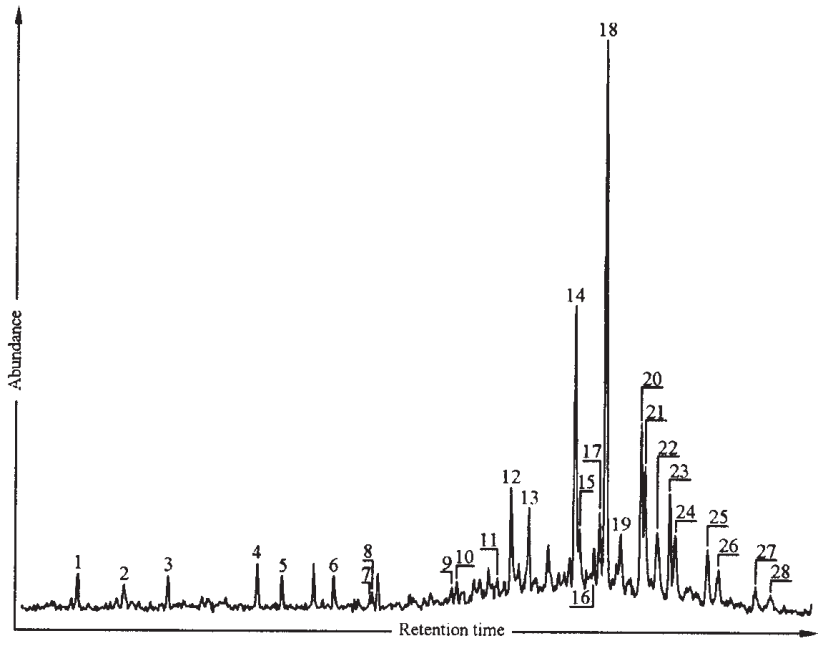


Fig. 5. GC-MS fragmentogram of terpanes (*m/z* 191) isolated from sample Velebit 3, typical for the investigated oils.

ment HRGC 5160 gas chromatograph, fitted with a SCE BPX-5 capillary column (30 m × 0.25 mm ID × 0.25 μm df was used; temperature program: 80 ºC, isothermal time 3 min, heating up to 300 ºC at a heating rate of 3 ºC/min), coupled to a ThermoQuest Trace MS. Helium was used as the carrier gas (flow rate 1.5 mL $min^{-1}$). *n*-Alkanes and isoprenoids were identified from *m/z* 71, steranes and diasteranes from *m/z* 217 and triterpanes from *m/z* 191 fragmentograms. The most relevant peaks were identified based on organic geochemical literature data,[18–20] or on the basis of total mass spectra, using two mass spectra data bases.[21,22] Examples of *m/z* 71, 217 and 191 fragmentograms, typical for the investigated oils, are shown in Figs. 3 – 5, respectively. Identification of the peaks are reviewed in Tables II and III.

TABLE II. Identification of peaks from Fig. 4

| Symbol | Compounds |
|---|---|
| 1 | $C_{27}$13β(H)17α(H)20(*S*)-diasterane |
| 2 | $C_{27}$13β(H)17α(H)20(*R*)-diasterane |
| 3 | $C_{27}$13α(H)17β(H)20(*S*)-diasterane |
| 4 | $C_{27}$13α(H)17β(H)20(*R*)-diasterane |
| 5a | $C_{28}$13β(H)17α(H)20(*S*)24(*S*)-diasterane |
| 5b | $C_{28}$13β(H)17α(H)20(*S*)24(*R*)-diasterane |
| 6 | $C_{28}$13β(H)17α(H)20(*R*)24(*R*+*S*)-diasterane |
| 7 | $C_{28}$13α(H)17β(H)20(*S*)-diasterane + $C_{27}$14α(H)17α(H)20(*S*)-sterane |
| 8 | $C_{29}$13β(H)17α(H)20(*S*)-diasterane + $C_{27}$14β(H)17β(H)20(*R*)-sterane |
| 9 | $C_{28}$13α(H)17β(H)20(*R*)-diasterane + $C_{27}$14β(H)17β(H)20(*S*)-sterane |
| 10 | $C_{27}$14α(H)17α(H)20(*R*)-sterane |
| 11 | $C_{29}$13β(H)17α(H)20(*R*)-diasterane |
| 12 | $C_{29}$13α(H)17β(H)20(*S*)-diasterane |
| 13 | $C_{28}$14α(H)17α(H)20(*S*)-sterane |
| 14 | $C_{29}$13α(H)17β(H)20(*R*)-diasterane + $C_{28}$14β(H)17β(H)20(*R*)-sterane |
| 15 | $C_{28}$14β(H)17β(H)20(*S*)-sterane |
| 16 | $C_{28}$14α(H)17α(H)20(*R*)-sterane |
| 17 | $C_{29}$14α(H)17α(H)20(*S*)-sterane |
| 18 | $C_{29}$14β(H)17β(H)20(*R*)-sterane |
| 19 | $C_{29}$14β(H)17β(H)20(*S*)-sterane |
| 20 | $C_{29}$14α(H)17α(H)20(*R*)-sterane |

TABLE III. Identification of peaks from Fig. 5

| Symbol | Compounds |
|---|---|
| 1 | $C_{19}$ tricyclic terpane |
| 2 | $C_{20}$ tricyclic terpane |
| 3 | $C_{21}$ tricyclic terpane |
| 4 | $C_{23}$ tricyclic terpane |

TABLE III. Continued

| Symbol | Compounds |
|---|---|
| 5 | $C_{24}$ tricyclic terpane |
| 6 | $C_{25}$ tricyclic terpane |
| 7 | $C_{26}$ tricyclic terpane (22*S*) |
| 8 | $C_{26}$ tricyclic terpane (22*R*) |
| 9 | $C_{28}$ tricyclic terpane (22*S*) |
| 10 | $C_{28}$ tricyclic terpane (22*R*) |
| 11 | $C_{29}$ tricyclic terpane (22*S*) + $C_{29}$ tricyclic terpane (22*R*) |
| 12 | $C_{27}$18α(H)-22, 29,30-trisnorneohopane, Ts |
| 13 | $C_{27}$17α(H)-22,29,30-trisnorhopane, Tm |
| 14 | $C_{29}$17α(H)21β(H)-hopane |
| 15 | $C_{29}$18α(H)-30-norneohopane |
| 16 | $C_{29}$17β(H)21α(H)-moretane |
| 17 | Oleanane |
| 18 | $C_{30}$17α(H)21β(H)-hopane |
| 19 | $C_{30}$17β(H)21α(H)-moretane |
| 20 | $C_{31}$17α(H)21β(H)22(*S*)-hopane |
| 21 | $C_{31}$17α(H)21β(H)22(*S*)-hopane |
| 22 | Gammacerane |
| 23 | $C_{32}$17α(H)21β(H)22(*S*)-hopane |
| 24 | $C_{32}$17α(H)21β(H)22(*R*)-hopane |
| 25 | $C_{33}$17α(H)21β(H)22(*S*)-hopane |
| 26 | $C_{33}$17α(H)21β(H)22(*R*)-hopane |
| 27 | $C_{34}$17α(H)21β(H)22(*S*)-hopane |
| 28 | $C_{34}$17α(H)21β(H)22(*R*)-hopane |

*Statistical methods*

*Factor analysis*. Factor analysis is a statistical method used for combining a large number of data into a considerably smaller number of factors, representing groups of initially mutually linearly dependent parameters containing the same amount of information as their constituent parameters.[7] The values of the coefficients preceding the parameters, marked as loadings, define the significance of a particular parameter in the characterization of an analyzed group of samples. The significance of a particular factor is defined by its characteristic value and percent of variance.[9,23] In order to determine the relationship between the parameters for the sake of classification of the samples, an interdependence diagram of two factors may be constructed.[24] In the case when the parameters are defining factors reflecting certain types of reaction characteristic for the investigated group of samples, the course of these processes and their mutual agreement can be proven by constructing corresponding correlation diagrams of these factors.[25]

*Cluster analysis*. This method is used for the classification either of samples or parameters. Two basic types of cluster analyses are known: K-Means and hierarchical types. For K-Means analysis, it is necessary to define the number of groups into which the samples/parameters are to be clas-

sified, while hierarchical cluster analysis enables the grouping of the samples or parameters without any previous classification. Cluster analysis is based on a matrix measuring the differences between each parameter of each sample. These differences are squared. By adding the individual matrices, a summed matrix is obtained. In the case when the values of the parameters are essentially different, they should preliminarily be standardized, in order that in the final matrix each parameter becomes an equal share. Based on the final matrix, a dendrogram is constructed, which involves all samples or parameters being classified into groups on the basis of all data taken into consideration.[10,26]

*Discriminant analysis*. Discriminant anlaysis serves primarily for checking the proposed classification based on a greater number of considered parameters. Similar to the factors in factor analysis, the discriminant functions represent a combination of linearly connected parameters, and their relationship with the constituting parameters is determined by a coefficient of canonical correlation. The importance of the function is defined by a characteristic value, *i.e.*, by variance percentage and significance. The samples may be classified by constructing a diagram of discriminant functions. Moreover, it is possible to check whether preliminary considerations led to a proper definition of the groups. Also, based on the values of the parameter loadings in the discriminant function, their importance may be determined for the characterization and classification of the examined group of samples.[8,11,27]

## RESULTS AND DISCUSSION

### *General characteristics of the investigated oils*

On the basis of the *m/z* 71 fragmentograms (Fig. 3), it was concluded that the investigated oils were in the stage of a moderate (I) degree of biodegradation.[28] *n*-Alkanes were absent in all samples, the dominant compounds being isoprenoids, with pristane (Pr) and phytane (Phyt) as the most abundant. All the investigated oils were characterized by a ratio Pr/Phyt < 1, and for 19 out of 25 samples, this ratio covered a very narrow range of 0.5–0.7 (Table IV), an observation which indicated steady, predominantly reduction conditions of organic matter sedimentation. Sterane and triterpane biomarkers were not affected by biodegradation (Figs. 4 and 5). The values of the most often used ratios of sterane and terpane source and maturation parameters are summarized in Tables IV and V.

TABLE IV. Source parameters calculated from the distributions and abundances of steranes and terpanes, and the values of the Pr/Phyt ratio

| No | $\%C_{27}\alpha\alpha R$ | $\%C_{28}\alpha\alpha R$ | $\%C_{29}\alpha\alpha R$ | $G\times100/C_{30}H$ | $O\times100/C_{30}H$ | $C_{29}H/C_{30}H$ | Pr/Phyt |
|---|---|---|---|---|---|---|---|
| 1 | 36.17 | 29.98 | 33.85 | 12.92 | 13.95 | 0.55 | 0.63 |
| 2 | 40.59 | 27.67 | 31.74 | 13.46 | 14.65 | 0.53 | 0.54 |
| 3 | 37.62 | 29.95 | 32.42 | 12.20 | 14.24 | 0.52 | 0.60 |
| 4 | 38.56 | 32.82 | 28.62 | 11.64 | 14.33 | 0.58 | 0.71 |
| 5 | 38.22 | 32.05 | 29.73 | 10.71 | 15.65 | 0.60 | 0.65 |
| 6 | 37.90 | 30.94 | 31.15 | 11.51 | 15.57 | 0.51 | 0.61 |
| 7 | 40.97 | 28.28 | 30.75 | 15.11 | 14.26 | 0.56 | 0.63 |
| 8 | 38.18 | 33.30 | 28.52 | 12.39 | 13.58 | 0.51 | 0.92 |
| 9 | 39.10 | 31.62 | 29.28 | 10.14 | 14.36 | 0.54 | 0.62 |
| 10 | 39.91 | 31.73 | 28.36 | 11.49 | 13.68 | 0.51 | 0.41 |
| 11 | 35.06 | 34.80 | 30.14 | 14.09 | 15.27 | 0.54 | 0.66 |

TABLE IV. Continued

| No | %$C_{27}\alpha\alpha R$ | %$C_{28}\alpha\alpha R$ | %$C_{29}\alpha\alpha R$ | G×100/$C_{30}$H | O×100/$C_{30}$H | $C_{29}$H/ $C_{30}$H | Pr/Phyt |
|---|---|---|---|---|---|---|---|
| 12 | 41.79 | 28.71 | 29.50 | 13.65 | 16.38 | 0.54 | 0.55 |
| 13 | 40.18 | 32.99 | 26.83 | 11.70 | 13.77 | 0.60 | 0.64 |
| 14 | 38.51 | 31.82 | 29.66 | 9.73 | 15.02 | 0.58 | 0.70 |
| 15 | 42.44 | 28.17 | 29.39 | 13.72 | 18.01 | 0.65 | 0.47 |
| 16 | 38.38 | 32.20 | 29.41 | 11.54 | 15.28 | 0.56 | 0.64 |
| 17 | 39.78 | 29.48 | 30.73 | 12.58 | 13.44 | 0.54 | 0.56 |
| 18 | 39.76 | 32.12 | 28.11 | 13.34 | 17.74 | 0.55 | 0.66 |
| 19 | 39.15 | 30.47 | 30.38 | 10.70 | 19.02 | 0.54 | 0.63 |
| 20 | 44.02 | 30.10 | 25.88 | 11.11 | 14.36 | 0.61 | 0.61 |
| 21 | 44.40 | 28.62 | 26.98 | 11.74 | 16.95 | 0.55 | 0.63 |
| 22 | 40.62 | 29.14 | 30.24 | 12.01 | 15.27 | 0.52 | 0.39 |
| 23 | 47.32 | 28.43 | 24.25 | 13.85 | 14.53 | 0.60 | 0.51 |
| 24 | 44.19 | 30.18 | 25.63 | 15.16 | 21.46 | 0.62 | 0.66 |
| 25 | 30.87 | 29.74 | 39.38 | 31.68 | 25.00 | 0.65 | 0.40 |

$C_{27}\alpha\alpha R$-$C_{27}$14α(H)17α(H)20*R*-sterane; $C_{28}\alpha\alpha R$-$C_{28}$14α(H)17α(H)*R*-sterane;
$C_{29}\alpha\alpha R$-$C_{27}$14α(H)17α(H)20*R*-sterane; G-gammacerane;
O-oleanane; $C_{30}$H-$C_{30}$17(α)21(β)-hopane; $C_{29}$H-$C_{29}$17(α)21(β)-hopane.

TABLE V. Maturation parameters calculated from the distributions and abundances of steranes and terpanes

| No | αα(*S*)/ αα(*S*)+ αα(*R*) | ββ(*R*)/ ββ(*R*)+ αα(*R*) | dia(*S*)/ dia(*S*)+ster | $C_{31}$(*S*)/ (*S*)+(*R*) | $C_{30}$M/ $C_{30}$H | $C_{29}$M/ $C_{29}$H | Ts/ Ts+Tm | $C_{29}$Ts/ $C_{29}$H |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.46 | 0.59 | 0.28 | 0.58 | 0.11 | 0.14 | 0.51 | 0.20 |
| 2 | 0.47 | 0.58 | 0.26 | 0.58 | 0.11 | 0.13 | 0.53 | 0.19 |
| 3 | 0.44 | 0.57 | 0.27 | 0.57 | 0.11 | 0.15 | 0.56 | 0.21 |
| 4 | 0.51 | 0.64 | 0.28 | 0.58 | 0.11 | 0.13 | 0.49 | 0.18 |
| 5 | 0.47 | 0.61 | 0.31 | 0.57 | 0.10 | 0.13 | 0.53 | 0.20 |
| 6 | 0.46 | 0.60 | 0.30 | 0.57 | 0.11 | 0.15 | 0.54 | 0.24 |
| 7 | 0.45 | 0.59 | 0.32 | 0.53 | 0.12 | 0.13 | 0.54 | 0.24 |
| 8 | 0.50 | 0.62 | 0.33 | 0.56 | 0.11 | 0.13 | 0.50 | 0.22 |
| 9 | 0.48 | 0.63 | 0.29 | 0.57 | 0.10 | 0.14 | 0.55 | 0.19 |
| 10 | 0.48 | 0.62 | 0.30 | 0.55 | 0.10 | 0.13 | 0.50 | 0.19 |
| 11 | 0.47 | 0.61 | 0.35 | 0.56 | 0.13 | 0.20 | 0.50 | 0.24 |
| 12 | 0.48 | 0.62 | 0.33 | 0.63 | 0.12 | 0.17 | 0.59 | 0.27 |
| 13 | 0.50 | 0.64 | 0.33 | 0.54 | 0.11 | 0.12 | 0.58 | 0.24 |
| 14 | 0.48 | 0.60 | 0.31 | 0.56 | 0.10 | 0.15 | 0.55 | 0.21 |
| 15 | 0.43 | 0.58 | 0.30 | 0.55 | 0.13 | 0.11 | 0.60 | 0.21 |

TABLE V. Continued

| No | αα(*S*)/ αα(*S*)+ αα(*R*) | ββ(*R*)/ ββ(*R*)+ αα(*R*) | dia(*S*)/ dia(*S*)+ster | $C_{31}$(*S*)/ (*S*)+(*R*) | $C_{30}$M/ $C_{30}$H | $C_{29}$M/ $C_{29}$H | Ts/ Ts+Tm | $C_{29}$Ts/ $C_{29}$H |
|---|---|---|---|---|---|---|---|---|
| 16 | 0.48 | 0.60 | 0.25 | 0.55 | 0.11 | 0.14 | 0.56 | 0.21 |
| 17 | 0.44 | 0.58 | 0.30 | 0.59 | 0.10 | 0.18 | 0.60 | 0.26 |
| 18 | 0.45 | 0.57 | 0.25 | 0.59 | 0.11 | 0.17 | 0.52 | 0.20 |
| 19 | 0.44 | 0.54 | 0.24 | 0.56 | 0.12 | 0.15 | 0.54 | 0.19 |
| 20 | 0.50 | 0.62 | 0.32 | 0.60 | 0.12 | 0.11 | 0.56 | 0.21 |
| 21 | 0.45 | 0.58 | 0.29 | 0.55 | 0.11 | 0.15 | 0.52 | 0.28 |
| 22 | 0.43 | 0.60 | 0.30 | 0.59 | 0.10 | 0.15 | 0.55 | 0.28 |
| 23 | 0.51 | 0.63 | 0.28 | 0.49 | 0.14 | 0.17 | 0.59 | 0.25 |
| 24 | 0.48 | 0.60 | 0.33 | 0.59 | 0.17 | 0.14 | 0.58 | 0.22 |
| 25 | 0.38 | 0.55 | 0.39 | 0.52 | 0.19 | 0.33 | 0.54 | 0.35 |
| E.V. | 0.52–0.55 | 0.67–0.71 | / | 0.57–0.62 | 0.05–0.15 | / | / | / |

αα(*R*)-$C_{29}$14α(H)17α(H)20(*R*)-sterane; αα(*S*)-$C_{29}$14α(H)17α(H)20(*S*)-sterane; ββ(*R*)-$C_{29}$14β(H)17β(H)20(*R*)-sterane; ster- $C_{27}$14α(H)17α(H)20(*R*)-sterane; dia(*S*)- $C_{27}$13β(H)17α(H)20(*S*)-diasterane; $C_{31}$22(*S*)-$C_{31}$17α(H)21β(H)22(*S*)-hopane; $C_{31}$22(*R*)-$C_{31}$17α(H)21β(H)22(*R*)-hopane; $C_{30}$H-$C_{30}$17α(H)21β(H)-hopane; $C_{30}$M-$C_{30}$17β(H)21α(H)-moretane; $C_{29}$H-$C_{29}$17α(H)21β(H)-hopane; $C_{29}$M-$C_{29}$17β(H)21α(H)-moretane; Ts-$C_{27}$18α(H),22,29,30-trisnorneohopane; Tm-$C_{27}$17α(H),22,29,30-trisnorhopane; $C_{29}$Ts-$C_{29}$18α(H),30-norhopane; E. V. – equilibrium value[3]

*Sterane and terpane source parameters*

The observed distribution of regular $C_{27}$–$C_{29}$αα(*R*)-steranes suggested that the Velebit oils were of miscellaneous origin (Fig. 4; Table IV). The fact that oleanane and gammacerane were identified in all samples corroborated this observation (Fig. 5; Table IV). All the investigated samples were characterized by a domination of $C_{27}$14α(H)17α(H)-steranes and relatively similar values of gammacerane and oleanane indexes, as well as $C_{29}$ to $C_{30}$ hopane ratio (9.73–15.16, 13.43–21.47 and 0.51–0.65, respectively; Table IV). The only exception was the sample Velebit 25, containing, as the most abundant, the $C_{29}$ regular sterane, as well as showing the highest values of gammacerane and oleanane indexes (25.00 and 31.68, Table IV). These facts distinguished this sample from all others. In conclusion, based on general consideration of sterane and terpane source parameters, the sample Velebit 25 was classified, as the only one, into a separate group.

*Sterane and terpane maturation parameters*

The triterpane maturation parameters, $C_{30}$M/$C_{30}$H and $C_{31}$(22*S*)/(22*S*)+(22*R*) hopanes, were found to be within the equilibrium range in 23 and 15 of the investigated oils, respectively. On the other hand, neither the $C_{29}$14β(H)17β(H)20(*R*)/ $C_{29}$14α(H)17α(H)20(*R*) + $C_{29}$14β(H)17β(H)20(*R*) nor the $C_{29}$14α(H)17α(H)20(*S*)/

$C_{29}14\alpha(H)17\alpha(H)20(R)$ + $C_{29}14\alpha(H)17\alpha(H)20(S)$ sterane parameter was found to be within the equilibrium range (Table V).[3] Based on these observations, it was presumed that the maturity of the source rocks of the examined crude oils corresponded to vitrinite reflectance values of $0.65 < Ro \leq 1$. Furthermore, the sterane and terpane maturation parameters suggested the Velebit oils were of medium to high maturity. The parameters Ts/Ts+Tm, dia/ster and $C_{29}Ts/C_{29}H$, being maturity indicators, on the one hand, but depending on the effect of sedimentary mineral catalysts, on the other, showed similar values for all the investigated oils, indicating a homogeneous lithology of the corresponding source rocks, since maturity transformations are considerably more intensive in source than in reservoir rocks, especially under such mild conditions. Additionally, the values of the aforementioned parameters indicated that the source rocks of the Velebit oils were not of a carbonate-type. The sample Velebit 25 differed by having the highest dia/ster and $C_{29}Ts/C_{29}H$ values. Moreover, this oil differed from all others by its extreme values of the $C_{30}M/C_{30}H$, $C_{29}M/C_{29}H$ and $C_{29}14\alpha(H)17\alpha(H)20(S)/C_{29}14\alpha(H)17\alpha(H)20(R)$ + $C_{29}14\alpha(H)17\alpha(H)20(S)$ parameters (Table V). Hence, based on the sterane and terpane maturation parameters as well, the sample Velebit 25 can clearly be separated from all other samples. A detailed classification of the other investigated oils required analysis by multivariate statistical methods.

*Classification according to source parameters*

For a more detailed classification of the oils, a factor analysis of the source parameters was carried out (programme SPSS 10.0 for Windows) using varimax rotation (Table VI). According to up-to-date organic geochemical investigations, this method has been shown to be the most convenient.[7,9,29] The first factor was characterized by high loadings of the sterane source parameters, and high proportions of $C_{27}$ and $C_{29}$ regular steranes. The second factor was defined by the terpane parameters, *i.e.*, the $C_{29}$hopane/$C_{30}$hopane ratio, and gammacerane and oleanane indexes (bold numbers in Table VI). Again, in a diagram representing the ratio of factor 1 *vs.* factor 2 (Fig. 6a), only the sample Velebit 25 was clearly separated. All the other samples were grouped relatively close to each other and it was difficult to draw a sharp line between them. Because of this, the same set of five source parameters (proportion of $C_{27}\alpha\alpha R$ steranes, proportion of $C_{29}\alpha\alpha R$ steranes, $C_{29}H/C_{30}H$, $Gx100/C_{30}H$ and $Ox100/C_{30}H$) was subjected to cluster analysis using the Ward method, which according to up-to-date organic geochemical investigations was proven to be the most reliable. In the resulting dendrogram (Fig. 6b), not only was sample 25 (group I) clearly separated, but also samples 20, 21, 23 and 24 were separated forming a group II. Furthermore, the samples belonging to group III were actually indicating three separate subgroups: IIIa (samples 1, 3 and 11), IIIb (samples 2, 7, 12, 15, 17, 18 and 22), and IIIc (samples 4, 5, 6, 8, 9, 10, 13, 14, 16 and 19). Comparison of the results obtained by factor and cluster analyses (Fig. 6) showed good agreement, since both methods (the ratio of factor 1 *vs.* factor 2 as

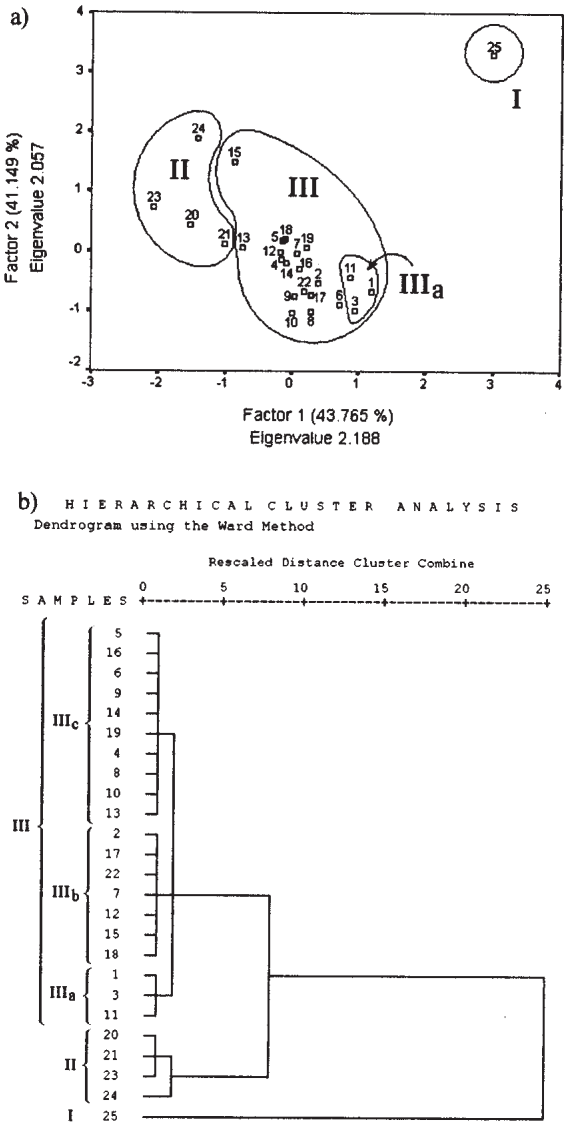


Fig. 6. Classification of the 25 oil samples according to sterane and terpane source parameters using factor (a) and Ward cluster analyses (b).

well as the Ward dendrogram) clearly distinguished the sample 25, and also grouped separately samples 20, 21, 23 and 24 (group II) and samples 1, 3 and 11 (group IIIa). Hence, it may be concluded that by using cluster analysis (Ward method), the samples are more precisely classified than by factor analysis.

In order to make a more detailed classification of the remaining 17 samples belonging to group III (Fig. 6), only partially separated by cluster analysis, the same set of source parameters was subjected to further examination. Considering the fact that the cluster analysis was shown to be more reliable, the 17 samples were first classified

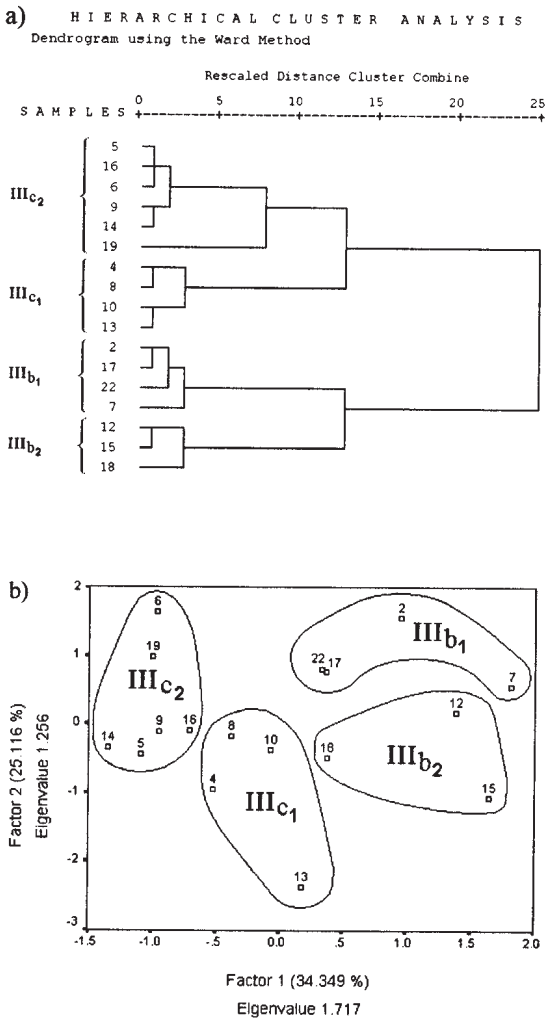


Fig. 7. Classification of the 17 oil samples belonging to group III, according to sterane and terpane source parameters, using factor (a) and Ward cluster analyses (b).

using this method (Fig. 7a). The observed results indicated the existence of two large groups corresponding to subgroups IIIb and IIIc, similar to previous analysis of the 25 samples (Fig. 6b). This observation confirmed the reproducibility and reliability of the Ward cluster method. Moreover, the dendrogram in Fig. 7a showed that the samples belonging to groups IIIb and IIIc can be further classified into separate subgroups: IIIb1 (2, 7, 17 and 22), IIIb2 (12, 15 and 18), IIIc1 (4, 8, 10 and 13), and IIIc2 (5, 6, 9, 14, 16 and 19). The seventeen oils belonging to group III were also classified by factor analysis (Fig. 7b). Comparing the results of cluster analysis (Fig. 7a) with the correlation diagram of factor 1 *vs.* factor 2, a good agreement was observed and the samples were likewise classified into two large groups, IIIb and IIIc, and four subgroups (IIIb1–IIIc2). Hence, by combining factor analysis with cluster analysis using the Ward method, 17 Velebit oil samples were classified, according to their origin, into

two groups, *i.e.*, into four subgroups (Fig. 7).

TABLE VI. Factor analysis of sterane and terpane source parameters using varimax rotation

| Parameter | Factors (% variance) | |
|---|---|---|
| | 1 (43.765) | 2 (41.149) |
| $\%C_{27}\alpha\alpha R$ | **–0.93** | $3.46 \times 10^{-3}$ |
| $\%C_{29}\alpha\alpha R$ | **0.95** | 0.16 |
| $G\times100/C_{30}H$ | 0.55 | **0.75** |
| $O\times100/C_{30}H$ | 0.26 | **0.85** |
| $C_{29}H/C_{30}H$ | –0.23 | **0.86** |

Kaiser–Meyer–Olkin measure of sampling adequacy (KMO) = 0.647; Significance = 0.000; Determinant = $4.787 \times 10^{-2}$

The classification of the 17 oils into four subgroups by Ward cluster analysis was checked by using K-Means cluster analysis. This method requires that the number of classification groups of the samples is determined in advance. The results of K-Means cluster analysis are shown in Table VII.

TABLE VII. Classification of 17 samples belonging to group III according to origin, using K-Means cluster analysis

| Sample | Subgroup |
|---|---|
| 2 | IIIb1 |
| 4 | IIIc1 |
| 5 | IIIc2 |
| 6 | IIIc2 |
| 7 | IIIb1 |
| 8 | IIIc1 |
| 9 | IIIc1 |
| 10 | IIIc1 |
| 12 | IIIb2 |
| 13 | IIIc1 |
| 14 | IIIc1 |
| 15 | IIIb2 |
| 16 | IIIc1 |
| 17 | IIIb1 |
| 18 | IIIb2 |
| 19 | IIIc2 |
| 22 | IIIb1 |

Comparing the results given in Fig. 7a and Table VII, only one difference was noticed, *i.e.*, with samples 9, 14 and 16. According to the Ward cluster analysis,

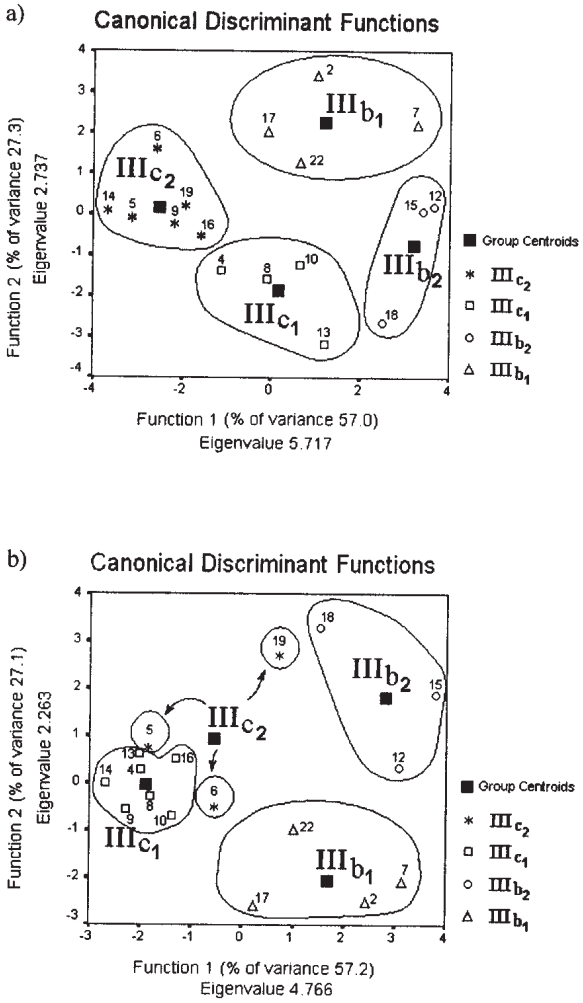


Fig. 8. Discriminant analysis of the 17 oil samples belonging to group III, according to origin, using Ward cluster (a) and K-means cluster analyses (b).

these samples were classified into subgroup IIIc2, whereas according to the K-Means cluster analysis, into subgroup IIIc1. In order to explain the observed discrepancy, a selection of source parameters characterizing the 17 investigated oils was subjected to discriminant analysis, the prerequisite of which is not only a definition of the number of groups but also the arrangement of the samples into groups. The results of discriminant analysis applied to categorizations, based on the Ward and K-Means cluster analyses are shown in Fig. 8. Identical classification of the samples into subgroups IIIb1 and IIIb2 as resulting from both the K-Means cluster and Ward cluster analyses, can be seen from the correlation diagrams of discriminant functions, which confirmed the accuracy of such a classification. However, in the case of oils belonging to subgroups IIIc1 and IIIc2, the K-Means cluster analysis yielded an unsatisfactory classification (Fig. 8b). On the other hand, the

discriminant analysis completely confirmed the classification of these samples previously assigned by the Ward method (Figs. 7a and 8a).

Hence, using factor, K-Means cluster, Ward cluster and discriminant analyses, the 25 oil samples were precisely classified into three groups according to their origin. The samples belonging to group III were further classified into five subgroups (Table VIII). Ward cluster analysis was shown to be the most reliable.

TABLE VIII. Classification of the oils according to sterane and terpane source parameters, based on the results of factor, cluster (K-Means and Ward) and discriminant analyses

| Group | Subgroup | Samples |
|---|---|---|
| I | / | 25 |
| II | / | 20, 21, 23 and 24 |
| III | a | 1, 3 and 11 |
| | b1 | 2, 7, 17 and 22 |
| | b2 | 12, 15 and 18 |
| | c1 | 4, 8, 10 and 13 |
| | c2 | 5, 6, 9, 14, 16 and 19 |

*Classification based on maturation parameters*

Classification of the oils according to maturation parameters was carried out in the same manner as the categorization based on source parameters. Considering the fact that so far the Ward cluster analysis was shown to be the most reliable, the 25 samples were first categorized on the basis of eight maturation parameters ($C_{30}M/C_{30}H$, $C_{29}M/C_{29}H$, $C_{31}(S)/(S) + (R)$, Ts/Ts+Tm, $C_{29}Ts/C_{29}H$, $C_{29}\beta\beta(R)/C_{29}\alpha\alpha(R) + C_{29}\beta\beta(R)$, $C_{29}\alpha\alpha(S)/C_{29}\alpha\alpha(R) + C_{29}\alpha\alpha(S)$ and dia/ster) using this statistical method. In the dendrogram shown in Fig. 9a, sample 25 was clearly separated (group I), confirming that this oil differed from all the other examined oils. According to thermal maturity, the remaining samples were categorized into two distinct groups: group II (1, 2, 3, 16, 18 and 19), and III (4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 20, 21, 22, 23 and 24). Furthermore, according to the previous results of the Ward cluster anlaysis, it was presumed that the samples belonging to group III could be classified into several subgroups. Factor analysis, using varimax rotation, of the same eight maturation parameters revealed 3 significant factors (Table IX). The first factor was characterized by high loadings in the case of the $C_{30}M/C_{30}H$, $C_{29}M/C_{29}H$, $C_{31}(S)/(S)+(R)$ and dia/ster parameters, the second was marked by the sterane maturation parameters $C_{29}\beta\beta(R)/C_{29}\alpha\alpha(R) + C_{29}\beta\beta(R)$ and $C_{29}\alpha\alpha(S)/C_{29}\alpha\alpha(R) + C_{29}\beta\beta(S)$, while the only significant parameter in the third factor was Ts/Tm + Ts (bold numbers in Table IX). This observation confirmed the results of previous investigations, *i.e.*, that the activation energy required for terpane isomerization reactions ($22R \rightarrow 22S$ and $\beta\alpha \rightarrow \alpha\beta$) are lower than those for sterane isomerization reactions ($22R \rightarrow 22S$ and $\beta\alpha \rightarrow \alpha\beta$) and that, hence, the terpane isomerization reactions occur at lower maturity lev-

els.[30,31] Classification of the parameters dia/ster and Ts/Ts+Tm, which otherwise depend on the effect of sedimentary mineral catalysts, into different factors is not unexpected since a disconcordance between these parameters has already been reported in some previous papers.[25,32]

TABLE IX. Factor analysis of sterane and terpane maturation parameters using varimax rotation

| Parameter | Factors (% variance) | | |
|---|---|---|---|
| | 1 (33.218) | 2 (28.395) | 3 (14.226) |
| ββ(*R*)/ββ(*R*)+αα(*R*) | $5.37 \times 10^{-2}$ | **0.97** | $-6.34 \times 10^{-3}$ |
| αα(*S*)/αα(*S*)+αα(*R*) | –0.24 | **0.91** | $-4.97 \times 10^{-2}$ |
| Dia(*S*)/dia(*S*)+ster | **0.85** | 0.21 | $6.26 \times 10^{-2}$ |
| $C_{31}$(*S*)/(*S*)+(*R*) | 0.45 | $8.19 \times 10^{-3}$ | 0.11 |
| $C_{29}M/C_{29}H$ | **0.74** | –0.52 | $-2.58 \times 10^{-2}$ |
| $C_{30}M/C_{30}H$ | **0.71** | –0.28 | 0.29 |
| Ts/Ts+Tm | $5.33 \times 10^{-2}$ | $-9.96 \times 10^{-3}$ | **0.98** |
| $C_{29}Ts/C_{29}H$ | 0.78 | –0.34 | 0.25 |

Kaiser–Mayer–Olkin measure of sampling adequacy (KMO) = 0.561; Significance = 0.000; Determinant = $7.746 \times 10^{-3}$

Similarly to the classification based on source parameters, an effort was made to classify the 25 Velebit oil samples according to thermal maturity using the ratio of factor 1 *vs*. factor 2 (Fig. 9b). However, in the corresponding correlation diagram, only sample 25 and a group of samples marked 2, 3, 16, 18 and 19 were clearly separated. When the results of Ward cluster analysis were applied to this correlation diagram, a very good concordance was observed and the samples were classified into three groups (Fig. 9). Again, the Ward analysis, compared to factor analysis, was shown to be more precise in fixing the boundaries between the groups.

In order to make a more detailed classification based on maturity, the remaining samples belonging to group III (Fig. 9) were further analyzed in the same way. Firstly, these 18 oils were classified applying the Ward cluster analysis on the same eight maturation parameters (Fig. 10a). The results suggested the existence of three subgroups of samples, which coincided with the subgroups resulting from the previous cluster analysis of the 25 samples (Fig. 9a), *i.e.*, IIIa (11, 12, 17, 21 and 22), IIIb (7, 13, 15, 20, 23 and 24), and IIIc (4, 5, 6, 8, 9, 10 and 14). This observation, again, confirmed the reliability of the method. Thereafter, a classification of the 18 oils belonging to group III was carried out by factor analysis using varimax rotation. However, in the correlation diagram of factor 1 *vs*. factor 2 (Fig. 10b) only samples belonging to subgroup IIIa (11, 12, 17, 21 and 22) were partly separated. Imprecise categorization based on factor analysis, in this case, was expected, taking into account that the sum of variance percentages for factors 1 and 2 was very low (45.60 %; Fig. 10b).

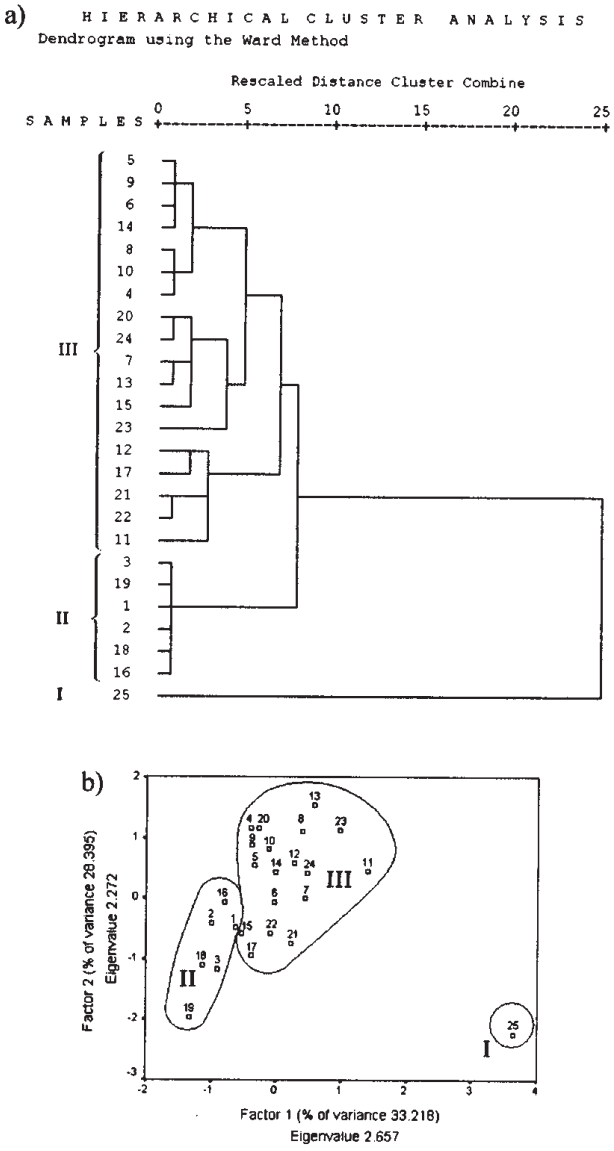


Fig. 9. Classification of the 25 oil samples according to sterane and terpane maturation parameters, using the Ward cluster analysis (a) and factor analysis (b).

Because of this, K-Means analysis was applied to the same eight maturation parameters bearing in mind the previously defined classification of the 18 samples into three groups according to their maturity. The results are listed in Table X. A good agreement was demonstrated between the K-Means and Ward cluster analyses (Fig. 10a and Table X) in the classification of the samples according to their maturity. The only discrepancies were observed with samples 6, 20 and 24.

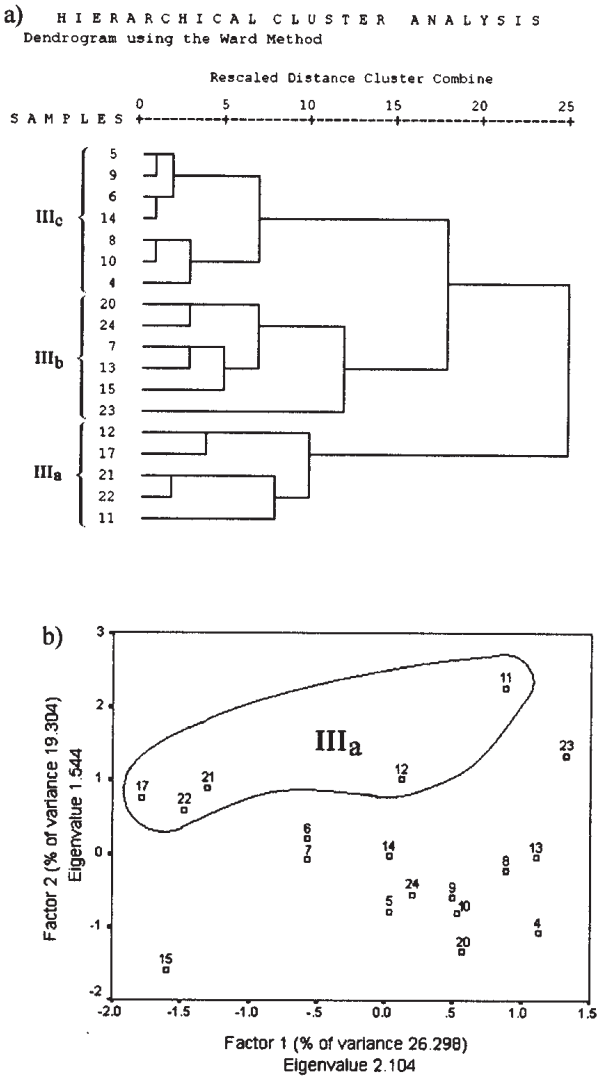


Fig. 10. Classification of the 18 oil samples belonging to group III, according to sterane and terpane maturation parameters, using Ward cluster analysis (a) and factor analysis (b).

In order to explain these discrepancies, categorizations based on Ward and K-Means cluster analyses were subjected to discriminant analysis. Correlation diagrams of the discriminant functions derived from the maturation parameters are given in Fig. 11. Samples 6 and 20 were more precisely classified by the K-Means method (Fig. 11b), in contrast to sample 24, which was more precisely classified by the Ward method (Fig. 11a).

TABLE X. Classification of the 18 samples belonging to group III according to maturity level, using K-Means cluster analysis

| Sample | Subgroup |
|---|---|
| 4 | IIIc |
| 5 | IIIc |
| 6 | IIIa |
| 7 | IIIb |
| 8 | IIIc |
| 9 | IIIc |
| 10 | IIIc |
| 11 | IIIa |
| 12 | IIIa |
| 13 | IIIb |
| 14 | IIIc |
| 15 | IIIb |
| 17 | IIIa |
| 20 | IIIc |
| 21 | IIIa |
| 22 | IIIa |
| 23 | IIIb |
| 24 | IIIa |

Hence, by combining four multivariate statistical methods, the 25 investigated oils were classified according to their thermal maturity into three groups. Samples belonging to group III were further categorized into three subgroups, IIIa – IIIc (Table XI).

TABLE XI. Classification of the oils according to sterane and terpane maturation parameters, based on the results of factor, cluster (K-Means and Ward) and discriminant analyses

| Group | Subgroup | Samples |
|---|---|---|
| I | / | 25 |
| II | / | 1, 2, 3, 16, 18 and 19 |
| III | a | 6, 11, 12, 17, 21 and 22 |
| | b | 7, 13, 15, 23 and 24 |
| | c | 4, 5, 8, 9, 10, 14 and 20 |

A general classification of the investigated oils into three main groups, according to source as well as maturation parameters, was possible only after detailed statistical analysis. The fact that these categorizations did not show any connection between individual groups of oils and the locations of the wells in the oil field suggests that these oils might have originated from more than one source rock, but from the same locality and of a similar generic type. Moreover, comparison of the classifications based on source and maturation parameters showed some samples to be in different groups (according to their origin and level of thermal maturity).

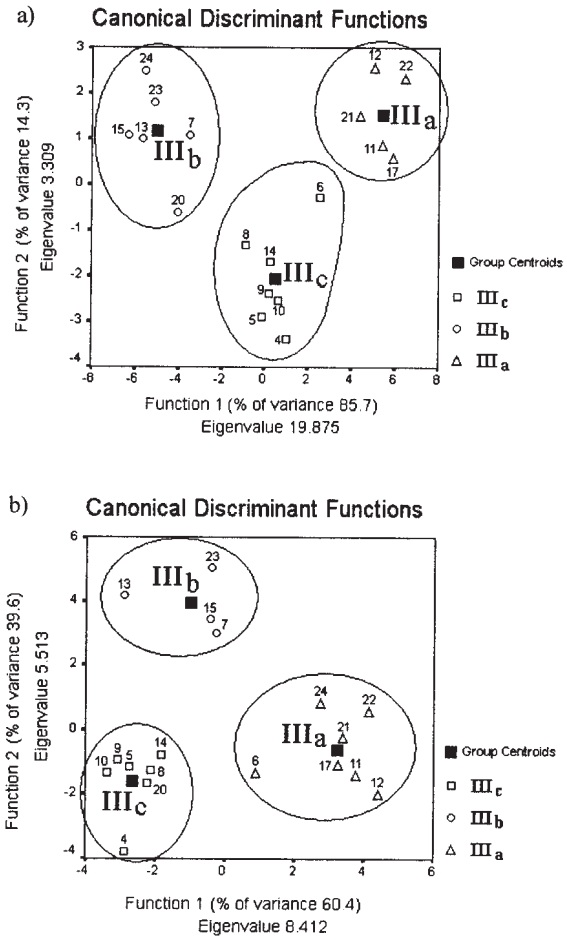


Fig. 11. Discriminant analysis of the 18 oil samples belonging to group III, according to maturity, using Ward cluster (a) and K-Means cluster analysis (b).

This fact implies possible mixing effects during the migration of the corresponding oils and indicates that the migration paths could not have been short.

## CONCLUSIONS

Saturated biomarkers were analysed in 25 crude oil samples originating from oil wells of the Velebit oil field. All samples were of moderate (I) degree of biodegradation, characterized by the absence of *n*-alkanes and undisturbed distributions of isoprenoids, steranes and triterpanes.

Source parameters, *i.e.*, the distribution of regular steranes, and oleanane and gammacerane indexes indicated miscellaneous origin, but of the same generic type, of the investigated oils. Based on the sterane and terpane maturation parameters, it was presumed that the relevant source rocks were of medium to high maturity, corresponding to vitrinite reflectance values of $0.65 < Ro \leq 1$. Moreover, the parameters dia/ster and neohopane/hopane indicated a homogeneous lithology of

the corresponding source rocks and that they could not have been of a carbonate-type.

In a general tabular presentation of the source and maturation parameters, only one sample was found to be separated into a distinctive group. Thereafter, a detailed classification of the samples based on source and maturation parameters was carried out using multivariate statistical methods: factor, cluster (K-Means and Ward hierarchical) and discriminant analyses, which allow the simultaneous examination of a large number of calculated parameters. By combining factor analysis with cluster analysis using the Ward method, the investigated oils were classified into three groups according to source and maturation parameters. The Ward method was shown to be the most reliable for the interpretation of saturated biomarkers in correlation studies of oils originating from an oil field in which all oil wells were at approximately the same depths of steady pressure and temperature.

A more precise classification of a large number of samples belonging to group III, into five subgroups according to origin, and into three subgroups according to level of thermal maturity, was feasible using factor, cluster (K-Means and Ward) and discriminant analyses. The discriminant analysis was shown to be the most significant for checking and confirmation of the proposed categorization.

Classification of the examined samples into three groups based on source and maturation parameters, was possible only after detailed statistical analysis. The fact that this categorization did not show any connection between the individual groups of oils and the locations of the wells in the oil field indicates the existence of a larger number of source rocks of these oils, but from the same locality and of a similar generic type. Moreover, the fact that some samples were located in different groups/subgroups (according to their source and level of thermal maturity) implies possible effects of mixing during their migration and indicated that migration path could not have been short.

## ИЗВОД

## МУЛТИВАРИЈАЦИОНЕ СТАТИСТИЧКЕ МЕТОДЕ У ИНТЕРПРЕТАЦИЈИ ЗАСИЋЕНИХ БИОМАРКЕРА (НАФТНО ПОЉЕ ВЕЛЕБИТ, ЈУГОИСТОЧНИ ДЕО ПАНОНСКОГ БАСЕНА, СРБИЈА)

ТАТЈАНА ШОЛЕВИЋ[1], КСЕНИЈА СТОЈАНОВИЋ[1,2], БРАНИМИР ЈОВАНЧИЋЕВИЋ[1,3], ГОРИЦА МАНДИЋ[4], JAN SCHWARZBAUER[5] и ДРАГОМИР ВИТОРОВИЋ[1,3]

*[1]ИХТМ – Центар за хемију, Његошева 12, 11000 Београд, [2]Медицински факултет, Универзитет у Београду, Вишеградска 26, 11000 Београд, [3]Хемијски факултет, Универзитет у Београду, п. пр. 158, 11001 Београд, [4]НИС – Нафтагас, Народног фронта 12, 21000 Нови Сад и [5]Institute of Geology and Geochemistry of Petroleum and Coal, Aachen University, Lochnerstrasse 4-20, D-52056, Aachen, Germany*

Испитивано је 25 сирових нафти нафтног поља Велебит (југоисточни део Панонског басена), најзначајнијег на територији Србије. Применом гаснохроматографско-масеноспектрометријске технике анализирани су биомаркери засићене фракције, *n*-алкани, изопреноиди, стерани и терпани. Израчунат је већи број изворних и матурационих

параметара, заснованих на расподели и обилности ових једињења, који се најчешће употребљавају у корелационим студијама нафти. Класификација узорака према пореклу и степену термичке зрелости изведена је применом факторске, кластерске и дискриминантне анализе. Факторска и кластерска анализа уз примену Ward методе омогућиле су категоризацију нафти у три групе према изворним и матурационим параметрима. При томе је кластерска Ward анализа показала већу осетљивост и поузданост. Међутим, за прецизнију и детаљнију категоризацију већег броја узорака из једне групе, поред две наведене методе била је неопходна и примена K-Means кластерске и дискриминантне анализе. Изведен је општи закључак да се за интерпретацију засићених биомаркера у корелационим студијама нафти могу употребљавати и факторска и кластерска K-Means и Ward анализа, али је добијене резултате неопходно проверити и потврдити дискриминантном анализом.

(Примљено 5. октобра 2005)

REFERENCES

1. A. S. Mackenzie, S. C. Brassell, G. Eglinton, J. R. Maxwell, *Science* **217** (1982) 491
2. D. W. Waples, T. Machihara, *Biomarkers for geologists – a practical guide to the application of steranes and triterpanes in petroleum geology*, The American Association of Petroleum Geologists, Oklahoma, 1991
3. K. E. Peters, J. M. Moldowan, *The Biomarker Guide, Interpreting Molecular Fossils in Petroleum and Ancient Sediment*, Prentice Hall, New Jersey, 1993
4. S. C. George, T. E. Ruble, A. Dutkiewicz, P. J. Eadington, *Appl. Geochem.* **16** (2001) 451
5. G. N. Gordzadze, *Thermolysis of organic substance in geochemical prospection explorations of oil and gas*, I. G. and R. G. I., Moscow, 2002 (in Russian)
6. W. K. Seifert, J. M. Moldowan, *Geochim. Cosmochim. Acta* **42** (1978) 77
7. K. G. Jöreskog, J. E. Klovan, R. A. Reyment, *Methods in Geomathematics I, Geological Factor Analysis*, Elsevier, Oxford, 1976
8. K. E. Peters, J. M. Moldowan, M. Schoell, W. B. Hempkins, in *Advances in Organic Geochemistry 1985*, Part 1, D. Leythaeuser, J. Rullkötter, Eds., Pergamon Press, Oxford, 1986, p. 17
9. C. Reimann, P. Filmoser, R. G. Garrett, *Appl. Geochem.* **17** (2002) 185
10. D. Globočanin, B. Škrbić, N. Miljević, *Chemom. Intell. Lab. Syst.* **72** (2004) 219
11. http://www.phychstat. smsu.edu/multibook/mlt03.htm
12. M. Ercegovac, *Petroleum geology*, RGF and DIT NIS-Naftagas, Belgrade, 2002 (in Serbian)
13. I. Csató, *Tectonophysics* **226** (1993) 377
14. F. Horváth, *Tectonophysics* **226** (1993) 333
15. C. S. Sajgó, Z. A. Horváth, J. Lefler, in *The Pannonian Basin – a study in basin evolution*, L. Royden, F. Horváth, Eds., AAPG Memoir 45, Tulsa-Oklahoma-Budapest, 1988, p. 297
16. F. Horváth, *Mar. Petrol. Geol.* **12** (1995) 837
17. S. Milosavljević, Dj. Marinović, S. Stanković, V. Aksin, *Results of oil and gas explorations in the Province of Vojvodina from 1949 to 1988 and perspectives of further explorations, DIT* **15-16** (1989) 9 (in Serbian)
18. P. R. Philp, *Fossil Fuel Biomarkers, Applications and Spectra*, Elsevier, Amsterdam, 1985
19. K. R. Al-Arouri, D. M. McKirdy, C. J. Boreham, in *Advances in Organic Geochemistry 1997*, Part 1, B. Horsfield, Ed., Pergamon Press, Oxford, 1998, p. 713
20. S. C. George, M. Lisk, R. E. Summons, A. R. Quezada, in *Advances in Organic Geochemistry 1997*, Part 1, B. Horsfield, Ed., Pergamon Press, Oxford, 1998, p. 631
21. NIST/EPA/NIH Mass Spectral Library NIST2000, Wiley
22. NBS Registry of Mass Spectral Data 7th Ed., electronic versions
23. M. D. Belonin, V. N. Golubeva, G. T. Skublov, *Principal Component Analysis in Geology*, Nauka, Moscow, 1982 (in Russian)

24. A. Golovko, G. Pevneva, in: *Abstracts of 21st International Meeting on Organic Geochemistry*, Krakow, September 8-12, 2003, GEOSPHERE, 2003, p. 73
25. A. Šajnović, K. Stojanović, I. Filipović, O. Cvetković, B. Jovančićević, D. Vitorović, *J. Serb. Chem. Soc.* **69** (2004) 611
26. http://www.psychstat.smsu.edu/multibook/mlt04.htm
27. D. Brodnjak-Vončina, Z. Cenčić-Kodba, M. Novič, *Chemom. Intell. Lab. Syst.* **75** (2005) 31
28. J. K. Volkman, R. Alexander, R. I. Kagi, G. W. Woodhouse, *Geochim. Cosmochim. Acta* **47** (1983) 785
29. A. Wilhelms, N. Telnæs, A. Steen, J. Augustson, *Org. Geochem.* **29** (1998) 97
30. K. A. Kvenvolden, J. P. Rapp, R. D. Hostettler, J. D. King, G. E. Claypool, in *Advances in Organic Geochemistry 1987*, L. Mattavelli, L. Novelli, Eds., Pergamon Press, Oxford, 1988, p. 351
31. J. Rullkötter, R. Marzi, in *Advances in Organic Geochemistry 1987*, L. Mattavelli, L. Novelli, Eds., Pergamon Press, Oxford, 1988, p. 639
32. A. G. Requejo, *Org. Geochem.* **21** (1994) 91.